

Estimation and Application of GARCH-X Model Based on High-Frequency Data

Zefang Song^{1*}, Lingjun Chen¹, Wenlin Huang²

¹School of Economics and Statistics, Guangzhou University, Guangzhou, China

²Department of Statistics, George Washington University, Washington, USA

Email: *song_zefang@163.com

How to cite this paper: Song, Z. F., Chen, L. J., & Huang, W. L. (2025). Estimation and Application of GARCH-X Model Based on High-Frequency Data. *American Journal of Industrial and Business Management*, 15, 242-259.

<https://doi.org/10.4236/ajibm.2025.152012>

Received: January 14, 2025

Accepted: February 17, 2025

Published: February 20, 2025

Copyright © 2025 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

This paper aims to study the GARCH-X model based on high-frequency data. Building upon the existing research on the selection criteria for optimal volatility representation and parameter estimation methods for high-frequency data GARCH(1,1) models, we extend the model to incorporate exogenous variables within high-frequency data GARCH(1,1) models and conduct both model simulations and empirical validation. The empirical research results show that in the study of the returns of the CSI 300 Index, including the 500 Index return rate as an exogenous variable can better explain the volatility of the return rate, thereby demonstrating the practical applicability of the model in real-world applications.

Keywords

High-Frequency Data, Exogenous Variables, GARCH-X Model, QMLE Estimation, VaR Forecasting

1. Introduction

Since the GARCH model was proposed by Bollerslev (1986), it has become one of the most commonly used models for describing and predicting market volatility in financial economics. It has been widely studied and applied by many researchers and financial institutions for risk management and resource allocation. With the continuous complexity and diversification of the financial market, various extensions of the GARCH model have emerged, such as the EGARCH model, GJR-GARCH model, GARCH-M model, and so on (see, e.g., Nelson, 1991; Engle et al., 1987; Glosten et al., 1993). Recent exploration in this area continues to thrive. For instance, Qian and Xu (2023) developed a novel closed-form option-pricing formula for general GARCH models. In the same vein, Wu, Zhao and Cheng (2023)

introduced the Real-Time GARCH-MIDAS model for volatility forecasting. These models enable the GARCH framework to better capture and predict volatility. Extensive theoretical and empirical research has been conducted to validate and apply these models. However, despite these valuable contributions, the understanding of GARCH, related models in financial markets remains incomplete.

From the perspective of incorporating exogenous variables, Han and Kristensen (2014) proposed the GARCH(1,1)-X model. They extended the traditional GARCH(1,1) model by directly adding an exogenous variable X to the conditional variance equation, quantifying the impact of the exogenous variable on volatility. Han (2015) also introduced the quasi-maximum likelihood estimation (QMLE) method for the GARCH(1,1)-X model and demonstrated its asymptotic properties under the conditions of the exogenous variables being stationary and ergodic. Lee (2017) extended the GARCH(1,1)-X model to the GARCH(p,q)-X model, further advancing the theoretical development of the GARCH-X model. Their research found that exogenous variables not only help explain the volatility of stock returns, exchange rates, and interest rates but also enhance the model's predictive ability. Singvejsakul et al. (2021) extrapolated the GARCH-X model based on Bayesian estimation methods and applied the model to analyze the fluctuation effects of various factors on cassava products in Thailand. Using the GARCH-X model, Apergis and Rezitis (2011) confirmed that food price fluctuations in Greece are positively correlated with short-term deviations from food prices and macroeconomic factors. These studies show that the GARCH-X model can provide implementations that the GARCH model and some other derivative models cannot capture, thus helping the financial market to obtain important information on time, and its emergence can also improve the financial theory more reasonably.

The development of big data storage technology, accompanied by increasingly convenient access to financial electronic transaction data, has made access to high-frequency data of intra-day trading a reality. Therefore, many scholars began to derive the GARCH model from low-frequency data to high-frequency data. The introduction of high-frequency data can provide more information, theoretically improve the asymptotic variance of parameter estimation, in practice, it can better reflect the fluctuation of the daily returns. However, high-frequency information contains too much microscopic noise, and the introduction of more information also means the introduction of more error terms, and the corresponding parameter estimation may have a large deviation.

This brings us to the main problem of introducing high-frequency data into models. Based on this, Visser (2009, 2011) proposed a volatility representative model, which uses a suitable volatility representative model to introduce intraday high-frequency information into the GARCH model so that the low-frequency GARCH model structure contains high-frequency information. Visser (2011) gives the quasi-maximum likelihood estimation of the GARCH model of high frequency and also provides the method of selecting the volatility representative model, indicating that the use of high-frequency data helps improve the estima-

tion accuracy of the GARCH model. Wang et al. (2018) provided parameter estimates for the high-frequency GARCH model using the composite quantile regression (CQR) method. Considering the advantages of realized variance in measuring intraday volatility and its strong empirical performance, Hansen et al. (2012) incorporated it as a new variable into the GARCH model to establish a new volatility model, known as the Realized GARCH model.

At the present stage, research on the GARCH-X model is mostly based on daily or monthly low-frequency data. The application of high-frequency data to the GARCH-X model is still insufficient. Considering the necessity of applying the GARCH-X model in financial market exploration and the effectiveness of high-frequency data in modeling and estimation, this paper combines high-frequency data with the GARCH-X model. This study aims to provide parameter estimation and optimal volatility representative selection based on high-frequency data within the GARCH-X model, thereby offering tools for capturing volatility in China's financial markets to better regulate the market.

The biggest improvement of this study compared to existing research lies in the consideration of exogenous variable volatility models while incorporating high-frequency data. This improvement allows for a more detailed capture of external factors affecting volatility, leading to more accurate volatility predictions and better risk management. The results of this study have verified this point.

The structure of the rest of this paper is as follows: Section 2 provides the construction of the GARCH-X model based on high-frequency data; Section 3 presents the QMLE estimation and theoretical properties of the model, as well as the selection of optimal volatility representatives; Section 4 conducts simulation analysis; Section 5 performs empirical analysis based on 1-minute high-frequency trading data from the CSI 300 index, verifying the volatility prediction ability of the new model and the accuracy of VaR prediction; Section 6 concludes the paper.

2. GARCH-X Model with High-Frequency Data

2.1. GARCH-X Model

Let $\{y_t\}_{t=1}^T$ be an observable return series. The GARCH(1,1)-X model with exogenous variables is constructed as follows:

$$\begin{cases} y_t = \sigma_t \varepsilon_t, & (1a) \\ \text{Var}(y_t | \mathcal{F}_{t-1}) = \sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma x_{t-1}^2 & (1b) \end{cases}$$

where $\{\varepsilon_t\}$ are independent and identically distributed random error terms with a mean of 0 and a variance of 1; when $t > s$, y_t and ε_s are independent. \mathcal{F}_{t-1} represents the low-frequency information set at time $t-1$, σ_t^2 is the conditional volatility, x_{t-1} represents the exogenous variable, and γ intuitively expresses the impact of the exogenous variable on the volatility. Set the parameters $\omega > 0$, $\alpha > 0$, $\beta \geq 0$, $\gamma > 0$ to ensure the non-negativity of the conditional variance. Since the GARCH(1,1) model is sufficient to characterize the ARCH effect in financial data, this paper uses the GARCH(1,1)-X model as the baseline model to

incorporate intraday high-frequency data.

2.2. Embedded Intraday Return Process of GARCH-X Model

To introduce high-frequency data, we consider extending the daily GARCH model using continuous-time processes. Let the intraday return process be denoted as $y_t(u)$, representing the logarithmic return of the asset at time u on day t . We standardize the time of each trading day by defining each trading day's time interval as $(0,1)$ and construct the following model:

$$\begin{cases} Y_t(u) = \sigma_t Z_t(u) & (2a) \\ \sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma x_{t-1}^2 & (2b) \end{cases}$$

Where $Z_t(\cdot)$ is a standard independent process. When $u=1$, $Y_t(1) = y_t$ and $Z_t(1) = \varepsilon_t$, the model is then a GARCH(1,1)-X model. Distinguished from low-frequency models, here σ_t is referred to as the scale parameter, and since $Z_t(\cdot)$ is independent, Visser (2011) refers to the above model as a scale model.

Note that model (2a-2b) contains high-frequency data $Y_t(u)$, $Z_t(u)$, and low-frequency data Y_{t-1}, x_{t-1} . Due to the inability to estimate parameters, based on the scale model, Visser (2011) further proposed a function that acts on intraday high-frequency data, referred to as the volatility representation. In general, the volatility representation is positive and satisfies positive homogeneity:

$H(pY(z)) = pH(Y(z))$. Utilizing this homogeneity, Visser cleverly embeds the volatility representation into low-frequency GARCH models, forming a class of volatility representation models that complete the embedding of high-frequency data. This paper adopts the same idea, embedding the volatility representation into the GARCH-X model.

From Homogeneity, we know:

$$H_t \equiv H(Y_t(u)) = H(\sigma_t Z_t(u)) = \sigma_t H(Z_t(u)), \quad (3)$$

Let

$$z_{H,t} = H(Z_t(u)), \mu = E(z_{H,t}^2), \varepsilon_t^* = \frac{z_{H,t}}{\sqrt{\mu}}. \quad (4)$$

Since $Z_t(u)$ is an independent and identically distributed sequence, $z_{H,t}, \varepsilon_t^*$ are also independent and identically distributed random variable sequences, and $E(\varepsilon_t^*) = 1$. Combining (2a-4), this paper constructs the following GARCH-X model to represent the embedded intraday return process:

$$\begin{cases} H_t = \sigma_t z_{H,t} = \sigma_t \sqrt{\mu} \varepsilon_t^*, & (5) \\ \sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma x_{t-1}^2. \end{cases}$$

3. Quasi-Maximum Likelihood Estimation of GARCH-X Model with High Frequency Data

3.1. Estimation of θ

We need to estimate parameter $\theta = (\omega, \alpha, \beta, \gamma, \mu)$ in the model, but because there

is an unknown redundant parameter μ in model(5), let $\sigma_t^* = \sigma_t \sqrt{\mu}$, then the model(5) can be transformed to

$$\begin{aligned} H_t &= \sigma_t z_{H,t} = \sigma_t^* \varepsilon_t^* \\ \sigma_t^{*2} &= \omega^* + \alpha^* y_{t-1}^2 + \beta^* \sigma_{t-1}^{*2} + \gamma^* x_{t-1}^2 \end{aligned} \tag{6}$$

where

$$\omega^* = \omega\mu, \alpha^* = \alpha\mu, \beta^* = \beta, \gamma^* = \gamma\mu. \tag{7}$$

The method of maximum likelihood estimation proposed by Visser (2011) allows us to first estimate $\theta^* = (\omega^*, \alpha^*, \beta^*, \gamma^*)'$ based on (6), then estimate μ , and lastly derive parameter from (7). According to the transformed model (6), the log quasi-likelihood function is expressed as:

$$L_T(\theta^*) = \sum_{t=1}^T l_t(\theta^*) = -\frac{1}{2} \sum_{t=1}^T \left(\ln(\sigma_t^{*2}(\theta^*)) + \frac{H_t^2}{\sigma_t^{*2}(\theta^*)} \right) \tag{8}$$

Thus, the quasi-maximum likelihood estimation (QMLE) for θ^* is defined as:

$$\hat{\theta}^* = \arg \max L_T(\theta^*). \tag{9}$$

At the same time, we can obtain $\hat{\sigma}_t^{*2} = \hat{\omega}^* + \hat{\alpha}^* y_{t-1}^2 + \hat{\beta}^* \sigma_{t-1}^{*2} + \hat{\gamma}^* x_{t-1}^2$. Next, we only need to obtain the estimate of μ to get the estimate of θ based on (7). It is noted that when the volatility representative $H_t = |y_t|$, model (6) corresponds to a low-frequency GARCH(1,1) model, the $\hat{\theta}^*$ derived through (8) and (9) corresponds to the parameter estimation of the model (1a-1b). Denoted as $\tilde{\theta} = (\tilde{\omega}, \tilde{\alpha}, \tilde{\beta}, \tilde{\gamma})^\tau$, thus obtaining the corresponding estimate sequence $\{\tilde{\sigma}_t^{*2}\}$. According to the aforementioned transformation, we find $\mu = \sigma_t^{*2} / \sigma_t^2$, thus the estimate for μ is:

$$\hat{\mu} = \frac{1}{T} \sum_{t=1}^T \frac{\hat{\sigma}_t^{*2}}{\hat{\sigma}_t^2}, \tag{10}$$

thus, we can obtain:

$$\hat{\theta} = (\hat{\omega}, \hat{\alpha}, \hat{\beta}, \hat{\gamma})^\tau = \left(\frac{\hat{\omega}^*}{\hat{\mu}}, \frac{\hat{\alpha}^*}{\hat{\mu}}, \hat{\beta}^*, \frac{\hat{\gamma}^*}{\hat{\mu}} \right)^\tau. \tag{11}$$

3.2. The Asymptotic Property of $\hat{\theta}^*$

Denote $\theta_0^* = (\omega_0^*, \alpha_0^*, \beta_0^*, \gamma_0^*)^\tau$ as the true value of the corresponding parameter.

If we assume:

- (A1) The parameter space of $\theta^* \in \Theta \subset (0, \infty)^4$, Θ is compact set;
- (A2) $E[\log(\alpha_0^* \varepsilon_t^{*2}) + \beta_0^*] < 0$;
- (A3) $E(\varepsilon_t^{*4}) < \infty$.

Given that the conditions are met, then

$$\sqrt{T}(\hat{\theta}^* - \theta_0^*) \xrightarrow{L} N(0, \Sigma^*), T \rightarrow \infty, \tag{12}$$

where $\Sigma^* = A_0^{-1}B_0A_0^{-1}$, $(A_0)_{i,j} = E\left(\frac{\partial^2 l_t(\theta_0^*)}{\partial\theta_i^* \partial\theta_j^*}\right)$, $(B_0)_{i,j} = E\left(\frac{\partial l_t(\theta_0^*)}{\partial\theta_i^*} \frac{\partial l_t(\theta_0^*)}{\partial\theta_j^*}\right)$.

Note: (i) According to the proofs by Visser (2011) and Straumann and Mikosch (2006), let $v_t(\theta^*) = \sigma_t^{*2}(\theta^*)$, then equation (8) can be rewritten as:

$$\begin{aligned} L_T(\theta^*) &= \sum_{t=1}^T l_t(\theta^*) = -\frac{1}{2} \sum_{t=1}^T \left(\ln(v_t(\theta^*)) + \frac{H_t^2}{v_t(\theta^*)} \right) \\ &= -\frac{1}{2} \sum_{t=1}^T \left(\ln(v_t(\theta^*)) + \frac{\sigma_t^* \varepsilon_t^*}{v_t(\theta^*)} \right). \end{aligned} \tag{13}$$

Since ε_t^* is independent of σ_t^* and $v_t(\theta^*)$, and satisfies $E(\varepsilon_t^*) = 1$, we define $L(\theta^*) = E(L_0(\theta^*))$, then we have:

$$\begin{aligned} L(\theta^*) &= -\frac{1}{2} E \left(\ln(v_0(\theta^*)) + \frac{\sigma_0^* \varepsilon_0^*}{v_0(\theta^*)} \right) 2(L(\theta^*) - L(\theta_0^*)) - 1 \\ &= E \left(\ln \left(\frac{v_0(\theta_0^*)}{v_0(\theta^*)} \right) - \frac{v_0(\theta_0^*)}{v_0(\theta^*)} \right) \end{aligned} \tag{14}$$

Given that $\ln(x) - x \leq 1$, with equality if and only if $x = 1$, therefore, to maximize $L(\theta^*)$, it must be that $v_0(\theta^*) = v_0(\theta_0^*)$, i.e., $\theta^* = \theta_0^*$. Hence, it follows that $\frac{L_T(\theta^*)}{T}$ converges uniformly to $L(\theta^*)$, and the parameter θ^* converges almost surely to θ_0^* .

(ii) The first-order partial derivative and the second-order partial derivative of $l_t(\theta^*)$ respectively are:

$$\begin{aligned} \frac{\partial l_t(\theta^*)}{\partial(\theta_i^*)} &= \frac{1}{\sigma_t^{*2}(\theta^*)} \frac{\partial \sigma_t^{*2}(\theta^*)}{\partial\theta_i} - H_t^2 \frac{1}{(\sigma_t^{*2}(\theta^*))^2} \frac{\partial \sigma_t^{*2}(\theta^*)}{\partial\theta_i} \\ &= \left(1 - \frac{H_t^2}{\sigma_t^{*2}(\theta^*)} \right) \frac{1}{(\sigma_t^{*2}(\theta^*))^2} \frac{\partial \sigma_t^{*2}(\theta^*)}{\partial\theta_i} \end{aligned} \tag{15}$$

$$\begin{aligned} \frac{\partial^2 l_t(\theta^*)}{\partial\theta_i^* \partial\theta_j^*} &= \frac{1}{\sigma_t^{*4}(\theta^*)} \left(\frac{2H_t^2}{\sigma_t^{*2}(\theta^*)} - 1 \right) \frac{\partial \sigma_t^{*2}(\theta^*)}{\partial\theta_i} \frac{\partial \sigma_t^{*2}(\theta^*)}{\partial\theta_j} \\ &\quad + (\sigma_t^{*2}(\theta^*) - H_t^2) \frac{1}{\sigma_t^{*4}(\theta^*)} \frac{\partial \sigma_t^{*2}(\theta^*)}{\partial\theta_i \partial\theta_j} \end{aligned} \tag{16}$$

When $\theta^* = \theta_0^*$, we get

$$(A_0)_{i,j} = E \left(\frac{\partial l_t(\theta_0^*)}{\partial(\theta_i^*) \partial(\theta_j^*)} \right) = E \left(\frac{1}{\sigma_t^{*4}(\theta_0^*)} \frac{\partial \sigma_t^{*2}(\theta_0^*)}{\partial\theta_i} \frac{\partial \sigma_t^{*2}(\theta_0^*)}{\partial\theta_j} \right) \tag{17}$$

Since $H_t = \sigma_t^* \varepsilon_t^*$ and σ_t^* and ε_t^* are independent, we have:

$$\begin{aligned}
 (B_0)_{i,j} &= E\left(\frac{\partial l_t(\theta_0^*)}{\partial(\theta_i^*)} \frac{\partial l_t(\theta_0^*)}{\partial(\theta_j^*)}\right) \\
 &= E(\varepsilon_t^{*2} - 1) E\left(\frac{1}{\sigma_t^{*4}(\theta^*)} \frac{\partial \sigma_t^{*2}(\theta^*)}{\partial \theta_i} \frac{\partial \sigma_t^{*2}(\theta^*)}{\partial \theta_j}\right)
 \end{aligned}
 \tag{18}$$

Thus, the asymptotic variance obtained is $\Sigma^* = A_0^{-1} B_0 A_0^{-1}$. Since $var(\varepsilon_t^*) = E(\varepsilon_t^{*2}) = 1$, we have $var(\varepsilon_t^{*2}) = E((\varepsilon_t^{*2} - 1)^2)$.

Let $G(\theta_0^*)_{i,j} = E\left(\frac{1}{\sigma_t^{*4}(\theta^*)} \frac{\partial \sigma_t^{*2}(\theta^*)}{\partial \theta_i} \frac{\partial \sigma_t^{*2}(\theta^*)}{\partial \theta_j}\right)$, then the asymptotic variance of $\hat{\theta}^*$ is

$$\Sigma^* = var(\varepsilon_t^{*2}) G^{-1}(\theta_0^*)$$

It is not difficult to see that the elements of the matrix $G(\theta^*)$ and $\sigma_t^{*2}(\theta^*)$ are both related to the parameters $\theta^* = (\omega^*, \alpha^*, \beta^*, \gamma^*)$ and the first-order partial derivatives of $l_t(\theta^*)$. From this, we can conclude that:

$$\begin{aligned}
 \sigma_t^{*2} &= \omega^* + \alpha^* y_{t-1}^2 + \beta^* \sigma_{t-1}^{*2} + \gamma^* x_{t-1}^2 \\
 &= \frac{\omega^*}{1-\beta^*} + \alpha^* \sum_{i=0}^{\infty} (\beta^*)^i y_{t-i-1}^2 + \gamma^* \sum_{i=0}^{\infty} (\beta^*)^i x_{t-i-1}^2,
 \end{aligned}
 \tag{19}$$

and

$$\begin{aligned}
 \frac{\partial \sigma_t^{*2}}{\partial \omega^*} &= \frac{1}{1-\beta^*}, \quad \frac{\partial \sigma_t^{*2}}{\partial \alpha^*} = \sum_{i=0}^{\infty} (\beta^*)^{i-1} y_{t-i}^2, \\
 \frac{\partial \sigma_t^{*2}}{\partial \beta^*} &= \sum_{i=0}^{\infty} (\beta^*)^{i-1} \sigma_{t-i}^{*2}, \quad \frac{\partial \sigma_t^{*2}}{\partial \gamma^*} = \sum_{i=0}^{\infty} (\beta^*)^{i-1} x_{t-i}^2.
 \end{aligned}
 \tag{20}$$

Let the asymptotic variances of the estimates $\hat{\omega}, \hat{\alpha}, \hat{\beta}, \hat{\gamma}$ be $\sigma_{\omega}^{*2}, \sigma_{\alpha}^{*2}, \sigma_{\beta}^{*2}, \sigma_{\gamma}^{*2}$ respectively. Based on the asymptotic properties of $\hat{\theta}^*$, the following results are obtained:

$$\begin{aligned}
 \sqrt{T}(\hat{\omega} - \omega_0) &\xrightarrow{L} N(0, \sigma_{\omega}^{*2} / \mu^2) \\
 \sqrt{T}(\hat{\alpha} - \alpha_0) &\xrightarrow{L} N(0, \sigma_{\alpha}^{*2} / \mu^2) \\
 \sqrt{T}(\hat{\beta} - \beta_0) &\xrightarrow{L} N(0, \sigma_{\beta}^{*2} / \mu^2) \\
 \sqrt{T}(\hat{\gamma} - \gamma_0) &\xrightarrow{L} N(0, \sigma_{\gamma}^{*2} / \mu^2)
 \end{aligned}
 \tag{21}$$

3.3. Selection of Optimal Volatility Representation

We can see from the estimates that the choice of volatility representation affects the efficiency of parameter estimates. Therefore, we need to provide criteria for selecting the volatility representation. Based on the criterion of minimizing the asymptotic variance of the estimates, we proposed the optimal volatility representation selection criterion function for the general quasi-maximum likelihood estimation:

$$MH = \frac{E(H_t^4)}{(EH_t^2)^2}. \quad (22)$$

Li and Zhang (2021) prove that the smaller the value of MH , the smaller the asymptotic variance of the parameter estimates, leading to more accurate estimates. This paper also seeks to select the optimal volatility representation by finding the minimum value of MH .

4. Simulation

This paper uses numerical simulations to evaluate the actual effects of parameter estimation. To generate the simulation data for the model (2a-2b), we refer to Visser's method. First, we simulate the standard stochastic process $Z(u)$ with its simulation process expressed as follows:

$$\begin{aligned} d\Gamma_t(u) &= -\delta(\Gamma_t(u) - \mu_\Gamma)du + \sigma_\Gamma dB_t^{(2)}(u) \\ dZ_t(u) &= \exp(\Gamma_t(u))dB_t^{(1)}(u), u \in [0, 1] \end{aligned}$$

Where $B_t^{(1)}$, $B_t^{(2)}$ are two independent Brownian motions, and $Z(0) = 0$, $\Gamma(0)$ is randomly generated from a stationary distribution $(\mu_\Gamma, \sigma_\Gamma^2)$. The intraday time is set to a frequency of 1 minute, dividing the time interval $[0, 1]$ into 240 small intervals corresponding to the 1-minute frequency of a typical trading day. The parameters are set as $\delta = 1/2$, $\sigma_\Gamma = 1/4$, $\mu_\Gamma = -1/16$, allowing us to simulate the discretized standard process $Z(u)$.

Considering the mechanisms for generating two sets of data, they are specified as follows:

Model (1):

$$\begin{aligned} Y_t(u) &= \sigma_t Z_t(u), \\ \sigma_t^2 &= 0.25 + 0.1y_{t-1}^2 + 0.6\sigma_{t-1}^2 + 0.01x_{t-1}^2 \end{aligned}$$

Model (2):

$$\begin{aligned} Y_t(u) &= \sigma_t Z_t(u), \\ \sigma_t^2 &= 0.1 + 0.3y_{t-1}^2 + 0.2\sigma_{t-1}^2 + 0.05x_{t-1}^2. \end{aligned}$$

Where the exogenous variables in Model 1 follow a t-distribution, while those in Model 2 follow a standard normal distribution.

In the parameter estimation process, based on Visser's research, realized volatility (RV) is chosen as the representation of volatility, calculated as follows:

$$RV_t(k) = \sqrt{\sum_{i=1}^m [Y_t(u_{i,k}) - Y_t(u_{(i-1),k})]^2}, \quad (23)$$

Where k represents sampling frequency, and m represents the number of returns in a day. In the simulation, sampling frequencies of 5 minutes, 15 minutes, and 30 minutes are chosen to represent realized volatility $RV5$, $RV15$, and $RV30$, respectively. To illustrate the superiority of high-frequency information,

an additional volatility representation $H_t = |y_t|$ is considered for comparison. The number of trading days in the simulation is set as $T = 500, 700, 900$ and repeated for a total of 1000 times.

Table 1 and **Table 2** present the parameter estimation results under different volatility representations and trading days. Among them, Bias represents the estimation bias, the estimated standard deviation, and the Mean.MH is the iterative mean of the criterion function value MH estimated based on (22). From the results in the tables, regardless of the assumptions made about the model parameters and error terms, the estimation method proposed in this paper can accurately estimate unknown parameters. Moreover, as the sample size increases, the sample bias, standard deviation, and MH mean for the same volatility representation decrease. For the same sample size, the bias, standard deviation, and MH mean estimated using the realized volatility RV as the volatility representation are smaller than when using $|y_t|$ as the volatility representation. It can also be observed that as the sampling frequency of RV increases, the MH mean decreases. The effectiveness of the volatility representation in estimation can be ranked as follows: $RV5 > RV15 > RV30 > |y_t|$, which indicates that incorporating high-frequency data into the volatility representation model significantly improves the parameter estimation results.

Table 1. The estimated results of Model 1.

| | | | $ y_t $ | $RV5$ | $RV15$ | $RV30$ | |
|----------------|----------------|----------------|---------|----------|---------|---------|--------|
| $T = 500$ | $\hat{\omega}$ | Bias | 0.0449 | -0.0001 | 0.0022 | 0.0053 | |
| | | SD | 0.1966 | 0.0633 | 0.0727 | 0.0886 | |
| | $\hat{\alpha}$ | Bias | 0.0048 | 0.0034 | 0.0037 | 0.0045 | |
| | | SD | 0.0572 | 0.0218 | 0.0242 | 0.0285 | |
| | $\hat{\beta}$ | Bias | -0.0677 | -0.00677 | -0.0108 | -0.0157 | |
| | | SD | 0.2507 | 0.0813 | 0.0952 | 0.1164 | |
| | $\hat{\gamma}$ | Bias | 0.0132 | 0.0019 | 0.0028 | 0.0037 | |
| | | SD | 0.0295 | 0.0102 | 0.0119 | 0.0137 | |
| | Mean.MH | | | 4.0051 | 1.4279 | 1.5477 | 1.7233 |
| | $T = 700$ | $\hat{\omega}$ | Bias | 0.0319 | 0.0001 | 0.0015 | 0.0058 |
| | | | SD | 0.1723 | 0.0526 | 0.0623 | 0.0759 |
| | | $\hat{\alpha}$ | Bias | 0.0048 | 0.0017 | 0.0019 | 0.0028 |
| SD | | | 0.0527 | 0.0171 | 0.0204 | 0.0236 | |
| $\hat{\beta}$ | | Bias | 0.0095 | -0.0040 | -0.0064 | -0.0127 | |
| | | SD | 0.2257 | 0.0675 | 0.0817 | 0.0988 | |
| $\hat{\gamma}$ | | Bias | -0.0487 | 0.0010 | 0.0018 | 0.0024 | |
| | | SD | 0.0254 | 0.0089 | 0.0103 | 0.0119 | |
| Mean.MH | | | 3.9818 | 1.4083 | 1.522 | 1.6933 | |

Continued

| | | | | | | |
|----------------|----------------|------|---------|---------|---------|---------|
| <i>T</i> = 900 | $\hat{\omega}$ | Bias | 0.0302 | 0.0000 | 0.0004 | 0.0014 |
| | | SD | 0.1595 | 0.0492 | 0.0562 | 0.0660 |
| | $\hat{\alpha}$ | Bias | 0.0021 | 0.0020 | 0.0024 | 0.0022 |
| | | SD | 0.0422 | 0.0153 | 0.0175 | 0.0203 |
| | $\hat{\beta}$ | Bias | -0.0426 | -0.0031 | -0.0040 | -0.0056 |
| | | SD | 0.2046 | 0.0623 | 0.0726 | 0.0854 |
| | $\hat{\gamma}$ | Bias | 0.0073 | 0.0005 | 0.0009 | 0.0017 |
| | | SD | 0.0227 | 0.0075 | 0.0084 | 0.0100 |
| | Mean.MH | | 3.9487 | 1.3975 | 1.5137 | 1.6821 |

Table 2. The estimated results of Model 2.

| | | $ y_t $ | <i>RV</i> 5 | <i>RV</i> 15 | <i>RV</i> 30 | | |
|----------------|----------------|----------------|-------------|--------------|--------------|---------|--------|
| <i>T</i> = 500 | $\hat{\omega}$ | Bias | 0.0321 | -0.0034 | -0.0011 | 0.0025 | |
| | | SD | 0.1550 | 0.0504 | 0.0588 | 0.0710 | |
| | $\hat{\alpha}$ | Bias | 0.0071 | 0.0028 | 0.0027 | 0.0029 | |
| | | SD | 0.0603 | 0.0207 | 0.0232 | 0.0273 | |
| | $\hat{\beta}$ | Bias | -0.0632 | -0.0039 | -0.0074 | -0.0130 | |
| | | SD | 0.2333 | 0.0740 | 0.0869 | 0.1060 | |
| | $\hat{\gamma}$ | Bias | 0.0137 | 0.0032 | 0.0038 | 0.0045 | |
| | | SD | 0.0244 | 0.0069 | 0.0077 | 0.0091 | |
| | Mean.MH | | 4.1522 | 1.4744 | 1.5951 | 1.7747 | |
| | <i>T</i> = 700 | $\hat{\omega}$ | Bias | 0.0221 | -0.0011 | 0.0003 | 0.0030 |
| | | | SD | 0.1383 | 0.0409 | 0.0498 | 0.0586 |
| | | $\hat{\alpha}$ | Bias | 0.0066 | 0.0031 | 0.0031 | 0.0040 |
| SD | | | 0.0512 | 0.0178 | 0.0203 | 0.0234 | |
| $\hat{\beta}$ | | Bias | -0.0460 | -0.0053 | -0.0075 | -0.0128 | |
| | | SD | 0.2067 | 0.0595 | 0.0721 | 0.0862 | |
| $\hat{\gamma}$ | | Bias | 0.0106 | 0.0025 | 0.0030 | 0.0039 | |
| | | SD | 0.0193 | 0.0062 | 0.0069 | 0.0081 | |
| Mean.MH | | 4.1288 | 1.4579 | 1.5768 | 1.7554 | | |
| <i>T</i> = 900 | | $\hat{\omega}$ | Bias | 0.0204 | -0.0008 | 0.0003 | 0.0017 |
| | | | SD | 0.1242 | 0.0369 | 0.0430 | 0.0510 |
| | | $\hat{\alpha}$ | Bias | 0.0065 | 0.0022 | 0.0024 | 0.0031 |
| | SD | | 0.0470 | 0.0156 | 0.0179 | 0.0206 | |
| | $\hat{\beta}$ | Bias | -0.0449 | -0.0061 | -0.0085 | -0.0115 | |
| | | SD | 0.1861 | 0.0537 | 0.0635 | 0.0754 | |
| | $\hat{\gamma}$ | Bias | 0.0096 | 0.0021 | 0.0027 | 0.0032 | |
| | | SD | 0.0169 | 0.0051 | 0.0060 | 0.0068 | |
| | Mean.MH | | 4.0352 | 1.4342 | 1.5542 | 1.7288 | |

5. Real Data Analysis

5.1. Data Description

To demonstrate that the high-frequency GARCH-X model constructed in this paper has better predictive performance, we establish a high-frequency GARCH-X model using the CSI 300 Index as the research object, with the CSI 500 Index selected as the exogenous variable. The data covers a period from January 3, 2019 to July 15, 2021, totaling 615 trading days, with a sampling frequency of one minute. All data are sourced from the Tongdaxin Financial Terminal.

The CSI 300 Index is selected in this paper because it is one of the earliest indices launched in China, comprising the 300 most representative securities with large scale and good liquidity from the Shanghai Stock Exchange and the Shenzhen Stock Exchange. This index serves as a price revelation function that characterizes fluctuations in stock prices and is a key indicator that best reflects the overall trend of the CSI market. The CSI 500 Index, on the other hand, is composed of the top 500 stocks in terms of total market capitalization after excluding the constituent stocks of the CSI 300 Index and the top 300 stocks in total market capitalization among all A-shares. This index is an important indicator reflecting the overall performance of small-cap stocks. There is a certain correlation between these two indices. Therefore, this paper studies the CSI 500 Index as an exogenous variable together with the CSI 300 Index.

The opening hours of China's stock exchanges are from 9:30 AM to 11:30 AM and from 1:00 PM to 3:00 PM, totaling 4 hours or 240 minutes. Therefore, all price sequence data consist of 1-minute closing prices, generating 240 data points per day. Data at a 5-minute frequency yields 48 data points per day, while data at a 10-minute frequency produces 24 data points per day, and data at a 30-minute frequency results in 8 data points per day.

The daily logarithmic rate of returns is calculated as follows:

$$Y_t(u) = [\ln g(P_t(u)) - \log(P_{t-1}(1))] * 100.$$

Based on the calculation method for realized volatility presented in (23), we use high-frequency data at 1-minute, 5-minute, 10-minute, and 30-minute intervals to calculate the realized volatility of intraday returns, denoted as RV1, RV5, RV10, and RV30, respectively. **Figure 1** presents the time series plots of the four realized volatilities for the CSI 300 Index. As can be seen from **Figure 1**, there is no apparent periodicity in the high-frequency data of the CSI 300 Index. According to the ADF test, these four realized volatilities are stationary series (with a p-value of 0.03 for RV1 and 0.01 for the others). Therefore, predictions can be made based on these series.

Figure 2 and **Figure 3** present the realized volatility series plot and Q-Q plot of the exogenous variable, the CSI 500 Index, respectively. From **Figure 2**, it can be observed that the exogenous variable does not exhibit any apparent periodicity. According to the ADF test, these four sequences are stationary (with a p-value of 0.03 for RV1 and 0.01 for the others), and they do not follow a normal distribution.

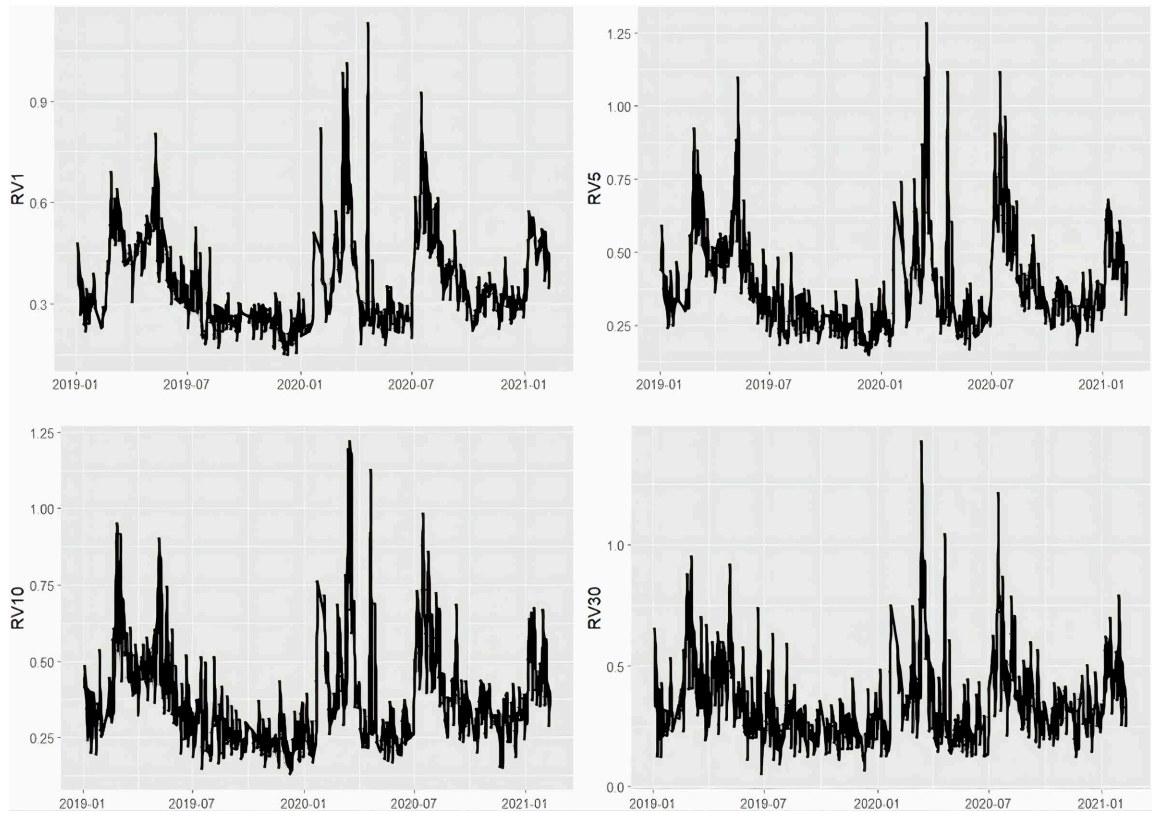


Figure 1. The realized volatility series plot of the CSI 300 Index.

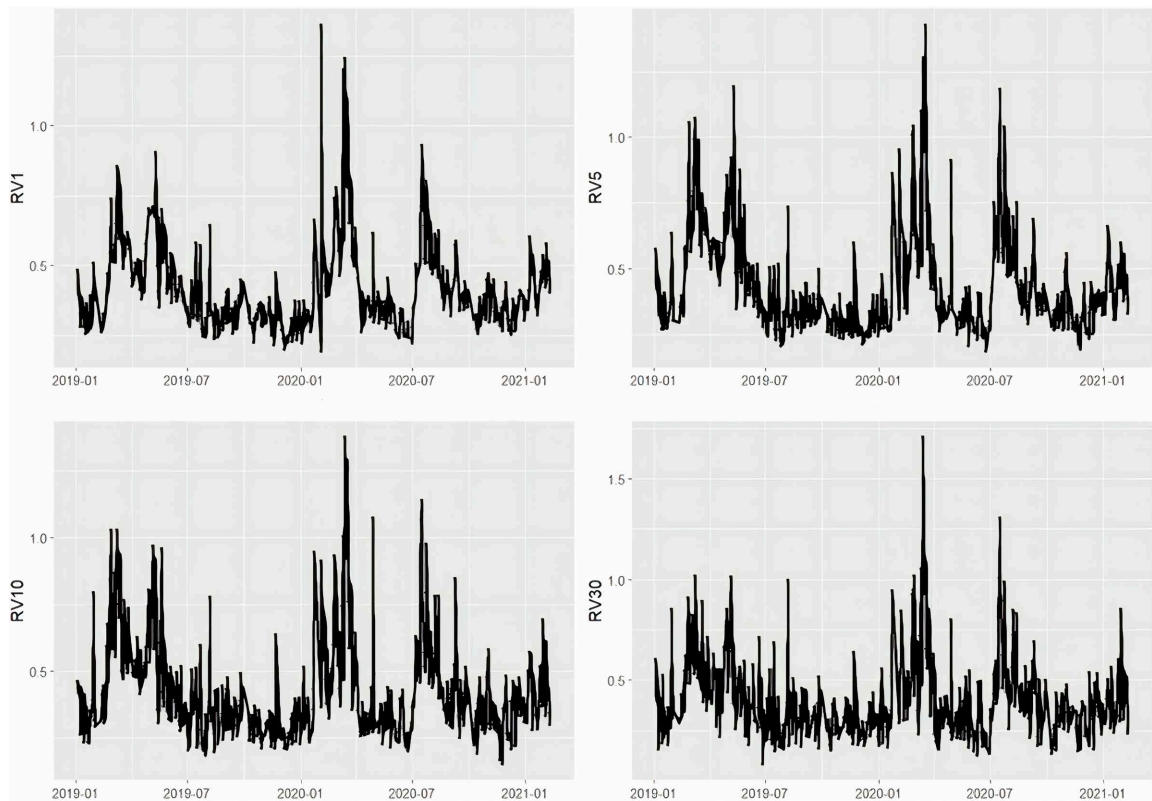


Figure 2. The realized volatility series plot of the CSI 500 Index.

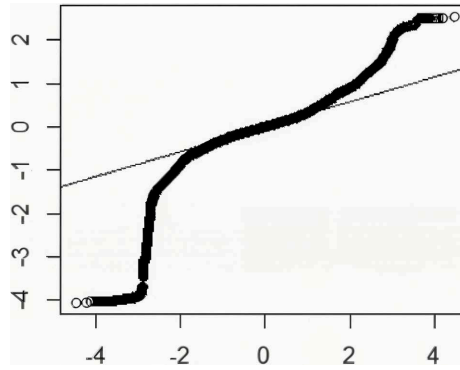


Figure 3. The Q-Q plot of the returns of the CSI 500 Index.

5.2. Estimation and Volatility Selection of High-Frequency GARCH-X Model

According to the aforementioned estimation method of the GARCH-X model under high-frequency data, we respectively give the parameter estimation results of the model and the corresponding MH estimates under different volatility representations, as shown in **Table 3**.

Table 3. Parameter estimation and MH value of High frequency GARCH-X model.

| H_t | $\hat{\omega}$ | $\hat{\alpha}$ | $\hat{\beta}$ | $\hat{\gamma}$ | \hat{MH} |
|---------|----------------|----------------|---------------|----------------|------------|
| $ y_t $ | 0.0187 | 0.1211 | 0.8007 | 0.0314 | 7.5758 |
| RV1 | 0.0044 | 0.2477 | 0.5402 | 0.14 | 1.8376 |
| RV5 | 0.0161 | 0.1065 | 0.6671 | 0.2951 | 2.1272 |
| RV10 | 0.0076 | 0.1103 | 0.8740 | 0.0059 | 2.1753 |
| RV30 | 0.0094 | 0.0947 | 0.8829 | 0.0018 | 2.4198 |

It can be seen from **Table 3** that among the four frequencies, the minimum MH value corresponds to the frequency of 1 minute. As the frequency decreases, the MH value increases, corresponding to the minimum $Var(\varepsilon_t^*)^2$ of the 1 minute. As the frequency decreases, the $Var(\varepsilon_t^*)^2$ value increases, and the MH value of all frequencies is smaller than that of the daily frequency. Therefore, it is appropriate to select the 1-minute volatility representation RV1 as the optimal volatility representation, and its corresponding estimation result as the final parameter estimation result.

Therefore, the results of the low-frequency GARCH(1,1)-X model fitting the CSI 300 index and selecting the 500 volatility index as exogenous variables are as follows:

$$\begin{aligned}
 y_t &= \sigma_t \varepsilon_t \\
 \hat{\sigma}_t^2 &= 0.0187 + 0.1211y_{t-1}^2 + 0.8007\hat{\sigma}_{t-1}^2 + 0.0314x_{t-1}^2
 \end{aligned}
 \tag{24}$$

The fitting result of the GARCH(1,1)-X model with realized volatility in 1 minute is as follows:

$$\begin{aligned}
 y_t &= \sigma_t \varepsilon_t \\
 \hat{\sigma}_t^2 &= 0.0044 + 0.2477 y_{t-1}^2 + 0.5402 \hat{\sigma}_{t-1}^2 + 0.14 x_{t-1}^2
 \end{aligned}
 \tag{25}$$

Based on the obtained fitting results, the volatility estimates based on daily and 1-minute realized volatilities are shown in **Figure 4**. It is evident from the figure that the trends of these two estimates are nearly identical.

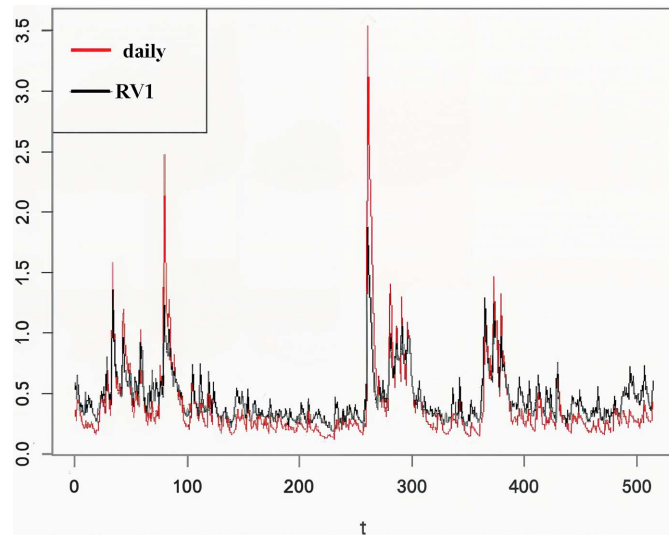


Figure 4. Volatility estimation plot.

5.3. VaR Estimation Using the High-Frequency GARCH-X Model

In today's financial markets, there is a strong emphasis on financial risk management. Value at Risk (VaR) stands out as an effective measurement method in risk management for financial markets, quantifying the potential maximum loss that a portfolio may face within a certain holding period and at a given confidence level. It has become a widely used metric.

This paper applies the estimation results of high-frequency data using the GARCH(1,1)-X model, namely (25), to the VaR estimation and forecasting of the CSI 300 Index. The VaR estimation process based on high-frequency data is as follows: Firstly, the parameters $\theta = (\omega, \alpha, \beta, \gamma)^\tau$ of the GARCH(1,1)-X model are estimated as $\hat{\theta} = (0.0044, 0.2477, 0.5402, 0.14)$ through model fitting using high-frequency data. Next, by substituting the estimated values $\hat{\theta}$ and daily logarithmic returns into the conditional variance formula of Equation (25) in the GARCH(1,1)-X model, we obtain 514 estimated conditional variances, denoted as \hat{h}_t .

Let $\varepsilon_t = \frac{H_t}{\sqrt{\hat{h}_t}}$, resulting in a sequence of ε_t . Setting the quantile $p = 0.1$, we take the 0.1 quantile of ε_t as $\hat{\varepsilon}_{0.1}$, finally, we set $\hat{q}_t = \hat{\varepsilon}_{0.1} \sqrt{\hat{h}_t}$, which represents the VaR estimate of H_t on day t at the $p = 0.1$ level at time $t - 1$.

For comparison and forecasting purposes, the first 514 trading days of data are used as the training set to estimate model parameters and VaR estimates. Then, a rolling forecasting method is employed to calculate the VaR forecast values for the

subsequent 101 samples (forecasting set). The VaR estimation results for the training set and the VaR forecasting results for the forecasting set are shown in **Figure 5**.

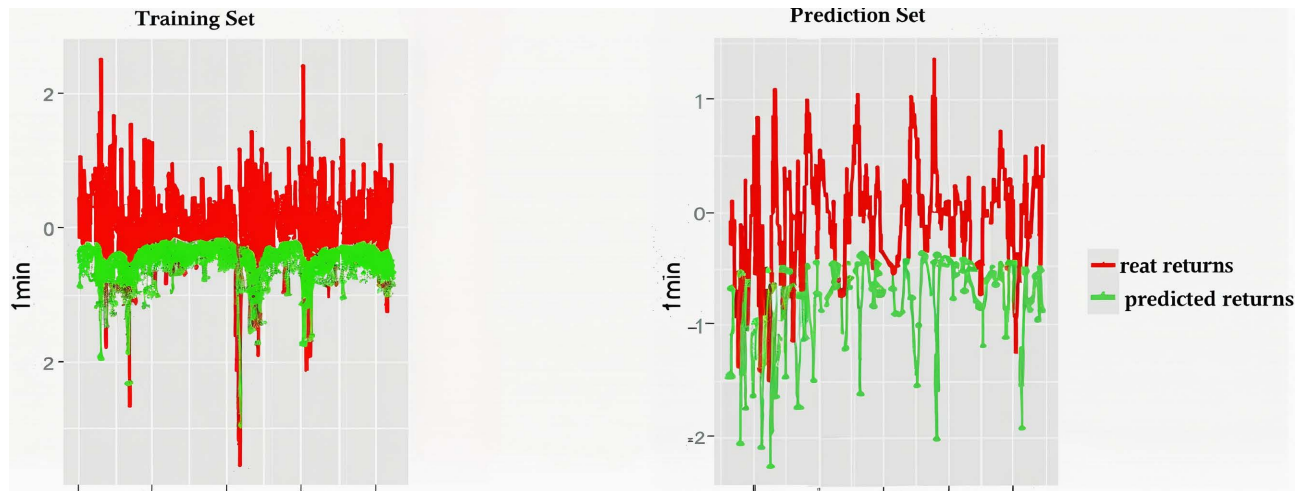


Figure 5. VaR variation chart for training and prediction sets of the CSI 300 Index.

The left graph in **Figure 5** shows the estimated results of VaR for the training set with a significance level of 0.1. It can be seen that there are over twenty instances where the predicted returns are greater than the actual returns during this period, while most of the time, the predicted returns are less than the actual returns. According to theory, there are 514 trading days in the training set, and 10% of the sample points, which amounts to an average of 52 trading days, should fall outside the VaR. This indicates that the prediction results for the training set are quite satisfactory. The right graph in **Figure 5** presents the prediction results for the forecasting set. The forecasting set consists of 101 trading days, and with a 0.1 quantile set, ideally, an average of 10 sample points fall outside the VaR. In the results of this forecasting set, there are 10 sample points that fall outside the VaR, which suggests that the prediction results for the forecasting set are also quite satisfactory. Overall, it can be concluded that the high-frequency GARCH model with exogenous variables has good predictive power for VaR.

To assess whether the high-frequency GARCH(1,1)-X model with exogenous variables can better reflect the estimation and prediction of the CSI 300 Index, the Mean Absolute Error (MAE) is introduced as a criterion for evaluating the VaR prediction results:

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|.$$

To test whether the high-frequency GARCH(1,1)-X model with exogenous variables can better reflect the estimation and prediction of the CSI 300 Index, referring to *Iqbal and Mukherjee's (2012)* test on the VaR prediction results of training set. Let $N^* = \sum_{i=1}^N I_i$, $\hat{p} = \frac{N^*}{N}$, where $I_i = I(H_i < \hat{q}_t)$, with N being

the total number of samples. Let $N_{i,j}$ be the count of points t such that $I_{t-1} = j$ and $I_t = i$ for $2 \leq t \leq N$, where $i, j = 0, 1$, also, let $\hat{\tau}_{i,j} = \frac{N_{i,j}}{N_{i,0} + N_{i,1}}$, $\hat{\tau} = \frac{N_{0,1} + N_{1,1}}{N}$. Define:

$$LR_{uc} = 2 \left\{ \log \left[\frac{(1-\hat{p})^{N-N^*} \hat{p}^{N^*}}{(1-p)^{N-N^*} p^{N^*}} \right] \right\},$$

$$LR_{ind} = 2 \left\{ \log \left[\frac{(1-\hat{\tau}_{0,1})^{N_{0,0}} \hat{\tau}_{0,1}^{N_{0,1}} (1-\hat{\tau}_{1,1})^{N_{1,0}} \hat{\tau}_{1,1}^{N_{1,1}}}{(1-\hat{\tau}^{N_{0,0}+N_{1,0}}) \hat{\tau}^{N_{0,1}+N_{1,1}}} \right] \right\},$$

then the statistic $LR_{cc} = LR_{uc} + LR_{ind}$ is asymptotically distributed as χ^2 .

Comparing the low-frequency GARCH model, high-frequency GARCH model, and high-frequency GARCH model with exogenous variables, the results are shown in **Table 4**.

Table 4. The comparison results of the models.

| Model | LRcc | MAE |
|---|--------|--------|
| Low-frequency GARCH model with exogenous variables | 23.5 | 0.6132 |
| High-frequency GARCH model with exogenous variables | 0.0255 | 0.4317 |
| High-frequency GARCH model | 3.15 | 0.4446 |

As can be seen from **Table 4**, in the training set, the LRcc value of the high-frequency GARCH model with exogenous variables is the smallest, followed by the GARCH model with high-frequency data, and the LRcc value of the GARCH model with low-frequency data is the largest. In the prediction set, the ordering of the MAE value is the same as the LRcc value ordering of the training set. This indicates that the prediction results of the high-frequency GARCH model with exogenous variables closer to the real value and that the prediction of the high-frequency GARCH model with exogenous variables is better and more effective.

6. Conclusion

Based on the traditional GARCH-X model, the QMLE estimation of the GARCH-X model with high-frequency data is introduced in this paper. According to the existing research on the volatility representation model and the parameter estimation method of the GARCH model, it is further extended to the parameter estimation research of the GARCH-X model and verified by simulation. In the empirical part, with 500 volatility return as an exogenous variable, the empirical analysis of the return of the CSI 300 index is as follows: 1) The optimal realized volatility is selected as RV1 by MH value; 2) High-frequency data can obtain more information. Through the estimation and prediction of VaR, it can be seen that the model with high-frequency data is better than the model with low-frequency

data, and the model with exogenous variables is better than the model without exogenous variables, which also verifies the validity of the high-frequency GARCH model with exogenous variables proposed in this paper and it can be better applied to the financial field.

This study can be generalized in several aspects. First, in this study, only realized volatility is introduced as the proxy for high-frequency data volatility, resulting in a relatively single variable selection. In fact, in existing research, there are numerous other volatility proxies that can be studied in depth, such as realized range, realized kernel estimation, etc. These different volatility proxies may capture different characteristics of financial market volatility, and are expected to bring new insights to the study of volatility in high-frequency data. Second, to further improve the model accuracy, more parameter estimation methods should be attempted in subsequent research, such as quasi-maximum exponential likelihood estimation. Lastly, this study mainly focuses on the basic GARCH model. In the future, it is advisable to consider expanding to other benchmark models such as GJR-GARCH, EGARCH, TGARCH, and FGARCH. In the context of high-frequency data, after introducing exogenous variables, compare the improvement effects of the predictive power of each model. Through comparative research, a more comprehensive understanding of the advantages and disadvantages of different models in high-frequency data prediction can be achieved, providing a more targeted basis for model selection in practical applications.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- Apergis, N., & Rezitis, A. (2011). Food Price Volatility and Macroeconomic Factors: Evidence from GARCH and GARCH-X Estimates. *Journal of Agricultural and Applied Economics*, 43, 95-110. <https://doi.org/10.1017/s107407080004077>
- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31, 307-327. [https://doi.org/10.1016/0304-4076\(86\)90063-1](https://doi.org/10.1016/0304-4076(86)90063-1)
- Engle, R. F., Lilien, D. M., & Robins, R. P. (1987). Estimating Time Varying Risk Premia in the Term Structure: The Arch-M Model. *Econometrica*, 55, 391-407. <https://doi.org/10.2307/1913242>
- Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *The Journal of Finance*, 48, 1779-1801. <https://doi.org/10.1111/j.1540-6261.1993.tb05128.x>
- Han, H. (2015). Asymptotic Properties of GARCH-X Processes. *Journal of Financial Econometrics*, 13, 188-221. <https://doi.org/10.1093/jjfinec/nbt023>
- Han, H., & Kristensen, D. (2014). Asymptotic Theory for the QMLE in GARCH-X Models with Stationary and Nonstationary Covariates. *Journal of Business & Economic Statistics*, 32, 416-429. <https://doi.org/10.1080/07350015.2014.897954>
- Hansen, P. R., Huang, Z., & Shek, H. H. (2012). Realized GARCH: A Joint Model for Returns and Realized Measures of Volatility. *Journal of Applied Econometrics*, 27, 877-906.

<https://doi.org/10.1002/jae.1234>

- Iqbal, F., & Mukherjee, K. (2012). A Study of Value-at-Risk Based on M-Estimators of the Conditional Heteroscedastic Models. *Journal of Forecasting*, *31*, 377-390. <https://doi.org/10.1002/for.1224>
- Lee, O. (2017). Some Limiting Properties for GARCH (p, q)-X Processes. *Journal of the Korean Data and Information Science Society*, *28*, 697-707.
- Li, L. L., & Zhang, X. F. (2021). Daily Frequency GARCH Model Estimation Based on High Frequency Data. *Journal of Guangxi Normal University (Natural Science Edition)*, *39*, 68-78.
- Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, *59*, 347-370. <https://doi.org/10.2307/2938260>
- Qian, Z., & Xu, X. (2023). An Option Valuation Formula for Stochastic Volatility Driven by GARCH Processes. *Journal of Mathematical Finance*, *13*, 221-247. <https://doi.org/10.4236/jmf.2023.132015>
- Singvejsakul, J., Chaovanapoonphol, Y., & Limnirankul, B. (2021). Modeling the Price Volatility of Cassava Chips in Thailand: Evidence from Bayesian GARCH-X Estimates. *Economies*, *9*, Article 132. <https://doi.org/10.3390/economies9030132>
- Straumann, D., & Mikosch, T. (2006). Quasi-Maximum-Likelihood Estimation in Conditionally Heteroscedastic Time Series: A Stochastic Recurrence Equations Approach. *The Annals of Statistics*, *34*, 2449-2495. <https://doi.org/10.1214/009053606000000803>
- Visser, M. P. (2009). *Volatility Proxies and GARCH Models*. University of Amsterdam.
- Visser, M. P. (2011). GARCH Parameter Estimation Using High-Frequency Data. *Journal of Financial Econometrics*, *9*, 162-197. <https://doi.org/10.1093/jffinec/nbq017>
- Wang, M., Chen, Z., & Wang, C. D. (2018). Composite Quantile Regression for GARCH Models Using High-Frequency Data. *Econometrics and Statistics*, *7*, 115-133. <https://doi.org/10.1016/j.ecosta.2016.11.004>
- Wu, X., Zhao, A., & Cheng, T. (2023). A Real-Time GARCH-MIDAS Model. *Finance Research Letters*, *56*, Article 104103. <https://doi.org/10.1016/j.frl.2023.104103>